

IN THE CLAIMS:

Please amend claims 1, 5, 9, 13, 14 and 15, and add new claims 16-21 as follows.

1. (Currently Amended) A method of speech recognition, said method comprising the steps of:

receiving audio signals from a speech source;

receiving video signals from the speech source;

detecting if the audio signals can be processed;

processing the audio signals if it is detected that the audio signals can be processed; ~~and- processing the video signals~~ if it is detected that at least a portion of the audio signal cannot be processed;

converting at least one of the audio signals and the video signals into recognizable information;

implementing a task based on the recognizable information.

2. (Original) The method of claim 1, wherein the step of receiving the video signals comprises the step of:

receiving video images of lip movements that coincide with the audio signals.

3. (Original) The method of claim 1, wherein the step of processing comprises the step of:

processing the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

4. (Original) The method of claim 1, further comprising the steps of:
storing the audio signals and the video signals; and
sending the audio signals and the video signals to a destination source.

5. (Currently Amended) A speech recognition device, said device comprising:
an audio signal receiver configured to receive audio signals from a speech source;
a video signal receiver configured to receive video signals from the speech source;
a processing unit configured to detect if the audio signals can be processed and if so, to process the audio signals and process the video signals if it is detected that at least a portion of the audio signals cannot be processed;

a conversion unit configured to convert at least one of the audio signals and the video signals to recognizable information;

an implementation unit configured to implement a task based on the recognizable information.

6. (Original) The device of claim 5, wherein the video signal receiver is configured to receive video images of lip movements that coincide with the audio signals.

7. (Original) The device of claim 5, wherein the processing unit is configured to process the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

8. (Original) The device of claim 5, further comprises:
a storage unit for storing the audio signals and the video signals; and
a transmitter for sending the audio signals and the video signals to a destination source.

9. (Currently Amended) A system for speech recognition, said system comprising:

a first receiving means for receiving audio signals from a speech source;
a second receiving means for receiving video signals from the speech source;
a processing means for detecting if the audio signals can be processed and processing the audio signals if the audio signals can be processed and for processing the video signals if at least a portion of the audio signals can not be processed;

a converting means for converting at least one of the audio signals and the video signals to recognizable information;

an implementing means for implementing a task based on the recognizable information.

10. (Original) The system of claim 9, wherein the second receiving means receives video images of lip movements that coincide with the audio signals.

11. (Original) The system of claim 9, wherein the processing means processes the audio signals and the video signals in parallel, wherein the video signals coincide with the audio signals.

12. (Original) The system of claim 9, further comprises:
a storage means for storing the audio signals and the video signals; and
a transmission means for sending the audio signals and the video signals to a destination source.

13. (Currently Amended) A method of speech recognition, said method comprising the steps of:

receiving audio signals from a speech source;

receiving video signals from the speech source;

detecting if the audio signals can be converted into a recognizable format;

processing the audio signals;

converting the audio signals into recognizable information;

processing the video signals when a segment of the audio signals can not be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;

converting the processed video signals into the recognizable information; and

implementing a task based on the recognizable information.

14. (Currently Amended) A speech recognition device, said device comprising:
an audio signal receiver configured to receive audio signals from a speech source;
a video signal receiver configured to receive video signals from the speech source;
a first processing unit configured to detect if the audio signals can be converted,
and if the audio signals can be converted, process the audio signals;

a first conversion unit configured to convert the audio signals to recognizable information;

a second processing unit configured to process the video signals when the audio signals cannot be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;

a second conversion unit configured to convert the processed video signals into the recognizable information; and

an implementation unit configured to implement a task based on the recognizable information.

15. (Currently Amended) A system for speech recognition, said system comprising:

a first receiving means for receiving audio signals from a speech source;

a second receiving means for receiving video signals from the speech source;

a first processing means for detecting if the audio signals can be converted, and if the audio signals can be converted, processing the audio signals;

a first converting means for converting the audio signals into recognizable information;

a second processing means for processing the video signals when a segment of the audio signals can not be converted into the recognizable information, wherein the video signals coincide with the segment of the audio signals that cannot be converted into the recognizable information;

a second converting means for converting the processed video signals into the recognizable information; and

an implementing means for implementing a task based on the recognizable information.

16. (New) The method of claim 1, wherein the step of detecting if the signals can be processed comprises:

defining an error threshold; and

comparing a number of errors detected in the audio signal with the threshold; and

determining that the audio signals can not be processed if the number of detected errors equals or exceeds the threshold.

17. (New) The speech recognition device according to claim 5, wherein the processing unit is configured to:

detect a number of errors in the audio signals;

compare the number of errors with a predefined threshold;

determining that the audio signals can not be processed if the number of detected errors equals or exceeds the threshold.

18. (New) The system for speech recognition according to claim 9, wherein the processing means is configured for:

detecting a number of errors in the audio signals;

comparing the number of errors with a predefined threshold; and

determining that the audio signals can not be processed if the number of detected errors equals or exceeds the threshold.

19. (New) The method of claim 1, further comprising:

determining if the video images of the user are detected; and

indicating to the user if the video image is not detected.

20. (New) The speech recognition device according to claim 5, wherein the processing unit is configured to:

determine if the user's video image is detected, and if the user's video image is not detected; and

indicate to the user that the video image is not detected.

21. (New) The system for speech recognition according to claim 9, wherein the processing means is further configured for:

determining if the user's video image is detected, and if the user's video image is not detected; and

indicating to the user that the video image is not detected.